Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR031308

Identification of Artificial Levees in the Contiguous United States

R. L. Knox[1], R. R. Morrison[2], and E. E. Wohl[1]

[1]Department of Geosciences, Colorado State University, Fort Collins, CO, USA, [2]Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, CO, USA

Key Points:

- An incomplete database of artificial levees inhibits the study of artificial levee impacts to floodplain extent in the contiguous USA
- Different methods of detecting artificial levees are tested in case study basins and in the contiguous USA
- This study detects 182,000 km of potential levees, along smaller streams in the Mississippi and Missouri Basins

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

R. L. Knox,
richknox@colostate.edu

Citation:

Knox, R. L., Morrison, R. R., & Wohl, E. E. (2022). Identification of artificial levees in the contiguous United States. *Water Resources Research*, *58*, e2021WR031308. https://doi.org/10.1029/2021WR031308

Received 27 SEP 2021
Accepted 11 APR 2022

Abstract Artificial levees are anthropogenic structures designed to hydrologically disconnect rivers from floodplains. The extent of artificial levees in the contiguous United States (CONUS) is unknown. To better estimate the distribution of artificial levees, we tested several different geomorphic, land cover, and spatial variables developed from the National Elevation Dataset, the National Land Cover database, and the National Hydrology Dataset HR Plus. We used known levee locations from the National Levee Database as training data. We tested machine learning and general logistic models' ability to detect artificial levees in a 100-year hydrogeomorphic floodplain of seven geographically diverse 8-digit HUC basins. Random forest models outperformed other models in predicting the location of levees using variables representing geomorphic attributes, land cover, and distance from streams ranging in size between stream order one through six. To demonstrate the ability of our approach to detect unknown levees, we conducted a leave-one-out cross-validation in the lower Mississippi Basin using approximately 1,100 artificial levees. This approach detected known levees constituting 94% of the total levee length in the basin. Scaling up to the CONUS, we applied a high performing (overall accuracy of 97%) random forest model using land cover and stream order variables. We detected 182,213 km of potential levees, mostly along streams of order 2–6 in the Mississippi and Missouri River Basins, indicating that the national levee database contains 20.4% of levee length. Potential levees and those documented in the national levee database modify 2% of the total length of streams in the contiguous United States.

Plain Language Summary There has been a lot of research exploring how humans have impacted river systems to include impacts from dams and roads at national and global scales. However, the study of artificial levees has been limited to local studies due to incomplete databases of levee location. Artificial levees are linear dirt mounds, built next to rivers to stop flooding of property built on floodplains. Because the location of every artificial levee is not known in the US, we do not know to what extent floodplains have been separated from their rivers. To ameliorate this, our study explores different methods to detect unknown levees and applies an algorithm that is 97% accurate to the contiguous U.S. Most levees are built within floodplains, so we limited our study to the 100-year floodplain. We tested different types of algorithms in a limited study with variables explaining levee shape, human land cover, and distance from streams of different sizes. Then, we applied the most effective models to the contiguous U.S. We located over 182,000 km of potential levees, potential areas with artificial levees not identified in the NLD, most of which were built on smaller streams and concentrated in the Mississippi and Missouri River basins.

1. Introduction

John Barry described the Mississippi River as pulsing "like the artery of the American heartland" (Barry, 2007). Recognized as the World's most engineered megariver (Knox & Latrubesse, 2016), the Mississippi River is emblematic of alterations made to United States (U.S.) rivers during the last three hundred years. With an estimated 98% of the nation's 5.3 million km of rivers impacted by human activities (Graf, 2001), it is difficult to understate the degree of human modification to U.S. rivers. Direct alterations take place within river corridors (which we define as the active channel[s] and floodplain) and include flow regulation, channel engineering (e.g., straightening or dredging), placer and aggregate mining, beaver trapping, floodplain draining, and levee construction (Wohl, 2018).

During recent decades, rivers have been increasingly appreciated as ecosystems worthy of preservation and restoration (Bunn et al., 2010; Castro & Thorne, 2019; Graf, 2001; Palmer et al., 2014). Understanding the importance

of river ecosystems across a broad spectrum of functions requires that we recognize the importance of longitudinal, lateral, and vertical connectivity to off-channel environments (Harvey & Gooseff, 2015; Kondolf et al., 2006; Ward, 1989). *Connectivity* describes the degree to which matter and organisms are able to move across areas in a landscape (Wohl, 2017). *Lateral connectivity*, which describes fluxes between the channel and adjacent riparian and floodplain landforms (Covino, 2017; Ward, 1989), reflects processes that create channel and floodplain topography and stratigraphy (Brooks, 2003), and thus aquatic and riparian habitats (Blanton & Marcus, 2009; Junk et al., 1989; Pennington et al., 2010; Ward et al., 1999).

Diverse human alterations of river form and processes either increase or limit connectivity in river ecosystems. Most studies on connectivity focus on the local scale (e.g., Briggs et al., 2013), although national-scale studies have assessed the effects of dams on longitudinal connectivity (e.g., Graf, 1999; Jones et al., 2019) and roads and railroads on lateral connectivity (e.g., Blanton & Marcus, 2009). One anthropogenic feature that adversely impacts lateral connectivity is *artificial levees*, which can be defined as raised linear features built between active channels and floodplains to contain peak flows in the channel (Tobin, 1995). Although artificial levees can strongly influence lateral connectivity within river corridors, their national-to global-scale effects have not been quantified in a manner similar to the effects of dams (e.g., Grill et al., 2019; Lehner et al., 2011).

Local and regional studies have found that artificial levees influence river hydrology by increasing stage at and upstream from levee locations and increasing downstream conveyance and flooding beyond the levees (Criss & Shock, 2001; Czech et al., 2016; Heine & Pinter, 2012; Tobin, 1995). Levees alter channel processes by inducing bed coarsening and incision caused by increased channel velocities (Frings et al., 2009). Levees also limit lateral connectivity and the exchange of nutrients, sediment, and organisms between the channel and floodplain, resulting in significant ecological harm (Blanton & Marcus, 2009; Sparks et al., 2017; Wohl, 2018). In the context of human societies, the presence of artificial levees can promote floodplain development and increase the vulnerability of populations and infrastructure to flood damage (Pinter, 2005; White et al., 2001). On the other hand, levees can be viewed as effective flood protection measures within their intended design standards and are uncomplicated and inexpensive to build (Tobin, 1995).

Artificial levee construction began in the U.S. in the early 1700s with landowners living alongside the lower Mississippi River (Wohl, 2005). By the 1800s, the majority of the basin was leveed with patchworks managed by individuals and communities in nascent levee boards (Hudson et al., 2008). Beyond the lower Mississippi River basin, the twentieth century became the era of federal levee construction with funding provided by the federal government and other entities (Wohl et al., 2017). Levee construction increased in scope well into the twentieth century, especially in the Midwestern and Eastern US, as the federal government and the US Army Corps of Engineers (USACE) became more involved in mitigating natural disasters (Wohl, 2005). Many levees were built in the Mississippi basin after the 1927 and 1937 floods and in California after flooding in 1907 and 1909 (ASCE, 2017). The nation's focus on artificial levees as a prime flood protection tool, resulting in billions of dollars expended to construct thousands of kilometers of levee, has been described as a "levee love affair" (Tobin, 1995).

The length of artificial levees in the U.S. is unknown but estimates range between 48,000 and 167,000 km, corresponding to coverage of roughly 1% and 3% of total estimated river km in the contiguous US (ASCE, 2017; Heine & Pinter, 2012). The USACE started a national levee inventory in 2006, which resulted in the National Levee Database (NLD). Each levee is annotated by a line representing the levee crest and varying amounts of metadata. The NLD is currently estimated to be 30% complete (ASCE, 2017), but a comprehensive evaluation of the NLD's thoroughness has not been completed (Wing et al., 2017). Consequently, there is no national-scale assessment of how artificial levees have altered lateral connectivity on U.S. rivers (Wohl, 2017) analogous to Graf's national-scale assessments of the effects of dams on river longitudinal connectivity (1999, 2001).

Because artificial levees are constructed by people and share common morphological characteristics, many levees should be recognizable and distinguishable by their shape (Brown et al., 2017). Recent advances over the last two decades in the availability of high-resolution topography have revolutionized the ability to study landscapes (Passalacqua et al., 2015). Accordingly, nearly every study on the identification of artificial levees has exclusively used topography or topographic-derived geomorphic variables with the exception of two studies that used spectral signatures (Steinfeld et al., 2013; Steinfeld & Kingsford, 2013). Identification of levees at regional scales has used maximum curvature, entropy, and residual topography from lidar digital terrain models (Sofia et al., 2014).

At the national scale, Wing et al. (2019) used geomorphic variables derived from the National Elevation Dataset (NED) to determine which geomorphic features are important to retain in a hydrodynamic flood model during DEM coarsening, although the study was not specifically intended to recognize artificial levees.

Researchers have used different methods of analyzing variables to predict artificial levee location, including logistic regression and image segmentation (Steinfeld et al., 2013), statistical analysis (Sofia et al., 2014), hillshade and wavelet analysis (Czuba et al., 2015) and visual inspection (Czuba et al., 2015; Steinfeld et al., 2013). Although not yet applied to identifying artificial levees, the modeling capabilities of machine learning techniques make them a suitable application in the geosciences (Lary et al., 2016). Machine learning techniques, which include decision trees, neural networks, and support vector machines, are especially effective at recognizing patterns in complex data or in scenarios where the underlying principles are poorly understood (Valentine & Kalnins, 2016). Random forest modeling is a supervised machine learning technique that builds decision trees and predicts category labels, is relatively fast, and can result in high predictive accuracy compared to other machine learning techniques (Breiman, 2001; Choi et al., 2020). The high accuracy of random forest models compared to other machine learning techniques has made random forest a popular choice for detecting surface and subsurface anthropogenic features as well as hydrological predictions (Cho et al., 2019; Deines et al., 2017). However, the validity of methods used to detect artificial levees in previous studies and machine learning techniques to detect artificial levees at the regional or national scale are unknown.

We aim to improve upon these previous studies by employing and testing different categories of data (i.e., geomorphic, land cover type, and distance from stream) and different types of models (general logistic models [GLM], random forest models [RF], and support vector machine models [SVM]) to the specific problem of identifying artificial levee locations. We generated a calibrated 100-year hydrogeomorphic floodplain as the extent of our analyses because this floodplain contains most known levee locations and at 10% of the contiguous U.S. (CONUS) area, reduces the computational requirements of the CONUS analysis. We use a case study of seven different 8-digit hydrologic unit code (HUC8) watersheds (Seaber et al., 1987), to test the effects of sample size, ratio of levee to non levee data, and model variables on model accuracy. We use known levee locations from the NLD to generate data from levee and non levee locations for training and validation. We conduct a leave-one-out cross-validation on 1,171 levees in the lower Mississippi Basin to understand how a highly accurate RF model detects undocumented levees. At the CONUS level, we generate two different sized datasets (n = 30,600 and 3,060,000) for training and validation and test the accuracy of a selection of models and variables based on the results of the case study. We apply the most accurate model (trained on data sampled from 2,142,000 locations across the CONUS) to the entire CONUS floodplain and analyze the results to determine the location, length, and stream order association of potential levees that are not identified by the NLD. Our primary objective is to estimate the locations, spatial distribution, and stream order association of artificial levees across the contiguous U.S., especially as they relate to the completeness of the NLD.

We organize our work as follows. First, we provide a site description and justification of the seven HUC8 watersheds we chose as case studies along with information about our modeling approach and the data we used in the case study, in the leave-one-out cross validation, and in the CONUS study. We then present our results using different accuracy measures. Finally, we discuss implications for understanding the completeness of the NLD and ideas for future research.

2. Materials and Methods

2.1. Overview

In general, our modeling included the following major steps. First, we chose seven diverse 8-digit HUC basins for a case study (Figure 1). We then created a CONUS-scale 100-year hydrogeomorphic floodplain using GFPLAIN at a 30 m resolution (Nardi et al., 2019) in ArcGIS Pro (ESRI Inc., 2020). We used this floodplain as the studies' geographic extent. For each case study basin, we created multi-layered raster files with each variable represented by a layer. We used R software (R Core Team, 2020) and an NLD shapefile to test different variables, sample sizes, machine learners, and GLMs. Then, using R and ArcGIS Pro we conducted a leave-one-out cross-validation using ∼1,100 artificial levees in the HUC2 lower Mississippi River basin (Figure 1) to determine how the model predicts undocumented levees. We then applied our approach to the entire CONUS (Figure 1) and generated two differently sized data sets using Google Earth Engine (GEE; Gorelick et al., 2017) and ArcGIS Pro.

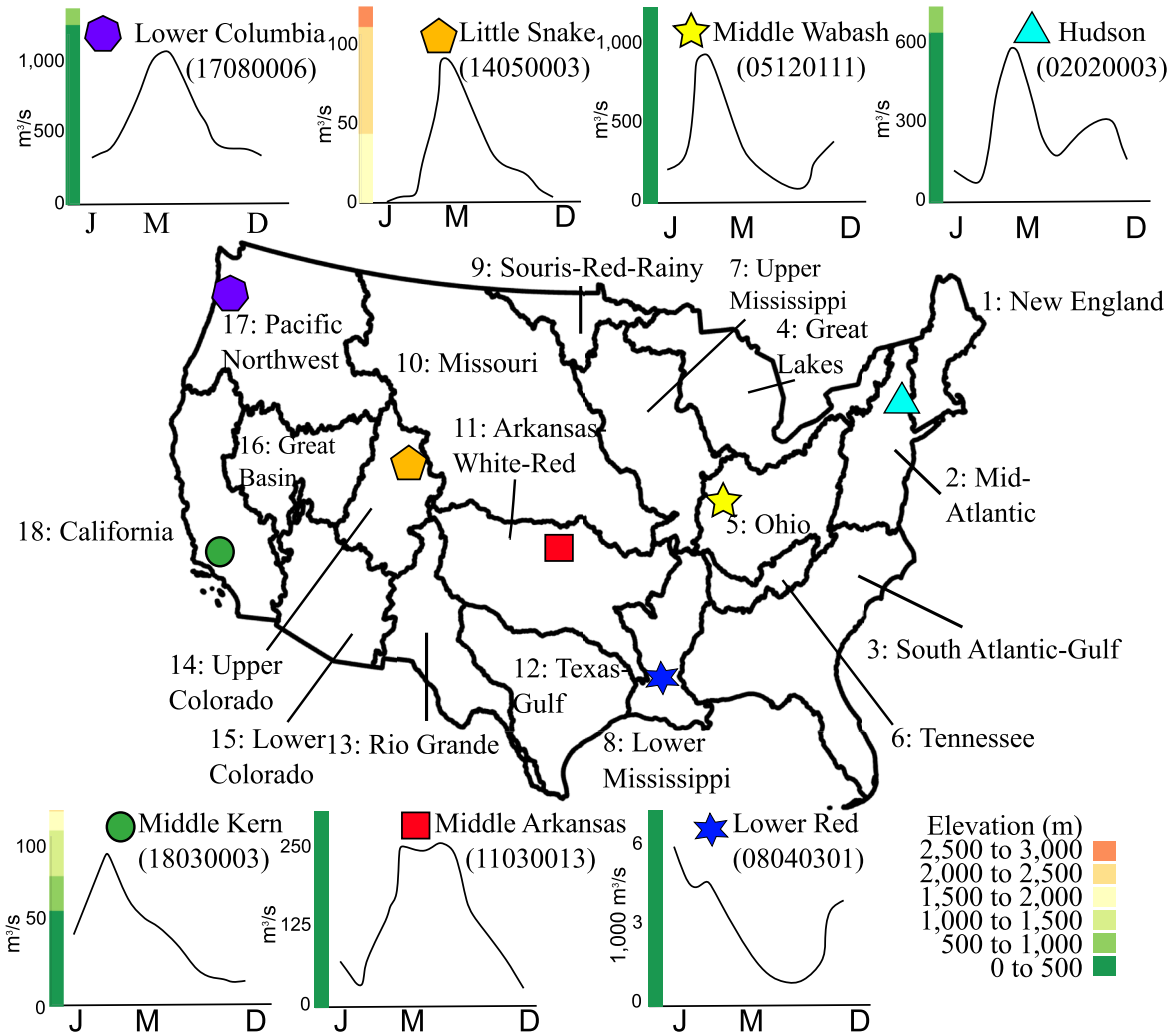

Figure 1. Location of the 18 HUC2 watersheds and seven HUC8 watersheds (HUC8 number listed below each basin's name) selected for the case study with average annual hydrographs and elevation bars (on the hydrographs' *y*-axes), representing the proportion of the basin in that elevation range.

Finally, we applied the most accurate model trained from the larger CONUS data set to the CONUS floodplain, creating a prediction surface map. This map was then segmented and compared to the NLD to determine which artificial levees were not documented.

Throughout, we used a full suite of accuracy metrics to assess model performance such as true and false rates and confusion matrices, when possible. In the case study, we used Cohen's kappa coefficient, which assesses inter-classifier agreement and is sensitive to class prevalence (Fitzgerald & Lees, 1994) to compare the performance of nearly 1,000 models. This coefficient, like all measures, is an imperfect index of overall accuracy (Foody, 2020). However, the comparison of many models using multiple measures of performance is unfeasible. To mitigate this, we test sample size, sampling ratio, model variables and type again in the national study, using multiple accuracy metrics to select which model to apply to the CONUS.

2.2. Site Descriptions

For the case studies, we chose seven distinct HUC8 watersheds across the 48 contiguous states to represent a wide variety of geographic, land cover, and hydrologic conditions (Figure 1, Table 1). The seven basins range in size from the 1,700 km^2 lower Columbia River in the Pacific Northwest to the 7,900 km^2 Little Snake River on the Colorado-Wyoming border. Climates of the seven locations range from the semi-arid Middle Kern River in southern California to the humid subtropical lower Red River in southern Louisiana and the humid continental

Table 1
Attributes of Each Basin Used in the Case Study

HUC8	Basin name	Primary land cover[a]	Maximum Strahler stream order[b]	Mean annual discharge (m^3/s)	Mean peak annual discharge (m^3/s)	Area (km^2)	Relief ratio	Principal sources of flood-causing precipitation or runoff[c]	References
17080006	Lower Columbia	Evergreen forest	9	7,533	14,915	1,754	0.017	Rain from extratropical cyclone on snowmelt	Simenstad et al. (2011); Cannon (2015)
02020003	Hudson	Deciduous forest	6	404	2,767	4,936	0.013	Rain from extratropical cyclone on snowmelt	Jackson et al. (2005)
18030003	Middle Kern	Herbaceous	6	46	149	6,779	0.031	Extratropical cyclone or associated front	Katibah (1984)
05120111	Middle Wabash	Cultivated crops	7	343	1,761	5,253	0.001	Rain from extratropical cyclone on snowmelt	Scheel et al. (2019)
14050003	Little Snake	Shrub/Scrub	6	32	147	7,926	0.001	Rain from extratropical cyclone on snowmelt	Blaschak (2012); Caskey (2013); Caskey et al. (2015)
08040301	Lower Red	Woody wetlands	8	6,173	15,282	2,350	0.001	Extratropical cyclone or associated front	Knox and Latrubesse (2016)
11030013	Middle Arkansas	Cultivated crops	7	124	800	2,350	0.002	Rain from extratropical cyclone on snowmelt	Guilliams (1998); Matthews et al. (2005)

[a]Source is the 2016 NLCD. [b]Stream order based on Strahler (1952). [c]Sources of flood-causing precipitation or runoff based on Hirschboeck (1991).

Hudson River in upstate New York. Relief ratios, the dimensionless ratio of the total vertical elevation difference in a basin divided by the basin length (Schumm, 1956), are an especially important consideration because artificial levees are distinguishable from other locations as particularly steep and rough terrain in the alluvial plains of the Midwest (e.g., 0.001 relief ratio of the Middle Wabash along the Illinois—Indiana border and the 0.002 relief ratio of the Middle Arkansas in Kansas) but seem more flat in comparison in the mountains of California or New York (e.g., 0.031 and 0.013 relief ratios of the Middle Kern and the Hudson, respectively). The primary land cover of each basin (Table 1) varies between cultivated crops (e.g., the Middle Wabash and the Middle Arkansas), to herbaceous or shrub/scrub (e.g., the Middle Kern and the Little Snake), to woody wetlands or forest (e.g., the Lower Red, the Hudson, the Columbia).

2.3. Hydrogeomorphic Floodplain Delineation

We established a CONUS floodplain calibrated to FEMA special flood hazard areas A and AE using the GFPLAIN algorithm (Nardi et al., 2019) and the 1 arc-second (~30 m resolution) NED (Table 2). The FEMA special flood hazard maps' coverage of the CONUS (~60% of the CONUS area) makes them suitable as validation for CONUS flood hazard models, but unsuitable as a stand-alone CONUS floodplain. Valued for its continental coverage, the accuracy of FEMA maps varies and can be less accurate than local high quality flood models (Blessing et al., 2017). Previous studies have used FEMA floodplain maps for either calibration or validation (e.g., Annis et al., 2022; Nardi et al., 2018; Wing et al., 2017), as the level of agreement between hydrogeomorphic models and other flood models indicates the suitability of this type of application, especially in data-poor areas (Lindersson et al., 2021). Lower model accuracy of hydrogeomorphic floodplains in certain areas (e.g., dry, steep, flat areas or those near the coast; Annis et al., 2022; Lindersson et al., 2021) is mitigated by our calibration at the 2-digit HUC basin level. The GFPLAIN algorithm identifies geomorphic floodplains in two main steps: (a) terrain analysis of a DEM for basin drainage extraction and (b) floodplain delineation. It uses an adaption of a scaling regression from Leopold and Maddock (1953) to relate stage to upstream contributing area:

$$FH_i = aA^b$$

where FH_i is the maximum flow depth at a location for the recurrence interval i, a, and b are dimensionless scaling parameters, and A is the contributing area for that location (Scheel et al., 2019). Calibration was conducted

Table 2
Variables and Data Sets Used in the Study

	Variable	Data Set	Type	Resolution
Geomorphic variables	Slope	National Elevation Dataset; Gesch et al. (2002)	Raster	10 m
	Planform curvature			
	Profile curvature			
	Relative elevation			
	Aspect difference			
Land cover variable	Land cover	National Land Cover Database, 2016; Jin et al. (2019)	Raster	30 m
Spatial variables	Distance from stream order 1 stream	National Hydrology Dataset Plus High Resolution; Buto and Anderson (2020)	Vector	–
	Distance from stream order 2 stream			
	Distance from stream order 3 stream			
	Distance from stream order 4 stream			
	Distance from stream order 5 stream			
	Distance from stream order 6 stream			
	Basin boundaries	USGS Hydrologic Unit Maps; Seaber et al. (1987)	Vector	–

by comparing model performance with FEMA flood maps on stream segments from the first six stream orders in each 2-digit HUC basin using the *F* measure of fit, which is the ratio of floodplain area correctly modeled to the total area modeled and predicted by FEMA (Equation 4 of Horritt & Bates, 2001). We kept parameter *a* constant at 0.0035 while varying *b* between 0.25 and 0.42, given the strong linear correlation of the two parameters (Annis et al., 2019). The range of *b* was selected based on previous studies (Annis et al., 2019; Nardi et al., 2006, 2018; Scheel et al., 2019). A third parameter, contributing area threshold, was kept constant at 50 km^2 based on previous studies (Annis et al., 2019; Scheel et al., 2019). Values for the *b* parameter 0.32, 0.34, and 0.36 resulted in the highest F measure of fit during floodplain calibration (Figure S1 in Supporting Information S1). The complete results are located in Supplemental Information.

2.4. Data

We used national, publicly available data sources (Table 2) in our analyses. The geomorphic variables (slope, planform curvature, profile curvature, relative elevation, and aspect difference) were developed according to Wing et al. (2019) from the National Elevation Dataset (NED 10 with resolution of 1/3 arc-second and approximately 10 m) based on its higher vertical accuracy compared to other nationally available topographic data sets (Gesch et al., 2014). We included the 2016 National Land Cover Database (NLCD) due to its national coverage and the expected association of artificial levees with certain land covers. We developed the six variables according to the distance from stream order one through six using the National Hydrology Dataset Plus High Resolution (NHD Plus) to capture the expected proximity of artificial levees to streams.

2.5. Model Development and Implementation

2.5.1. Case Study Modeling

Similar data processing and modeling procedures were followed in the case and national studies (Figure 2). For the case studies, R software was used to process data sources (R Core Team, 2020). Twelve raster layers (all of the variables in Table 2 except "basin," which was added after sampling), each with a 10 m resolution, were created for the calibrated 100-year hydrogeomorphic floodplains in each of the seven basins (Table 1). We randomly sampled levee locations from the NLD and non-levee locations from other places within the GFPLAIN floodplain using over 480 general logistic models (GLMs), random forest models (RFs), and support vector machine models (SVMs) with various sample sizes, non-levee to levee sample size ratios (absence/presence), and different combinations of variables (Table 3). We defined and tested the full model as model 1 with all variables: the five geomorphic variables (slope, profile curvature, planform curvature, relative elevation, and aspect difference), the

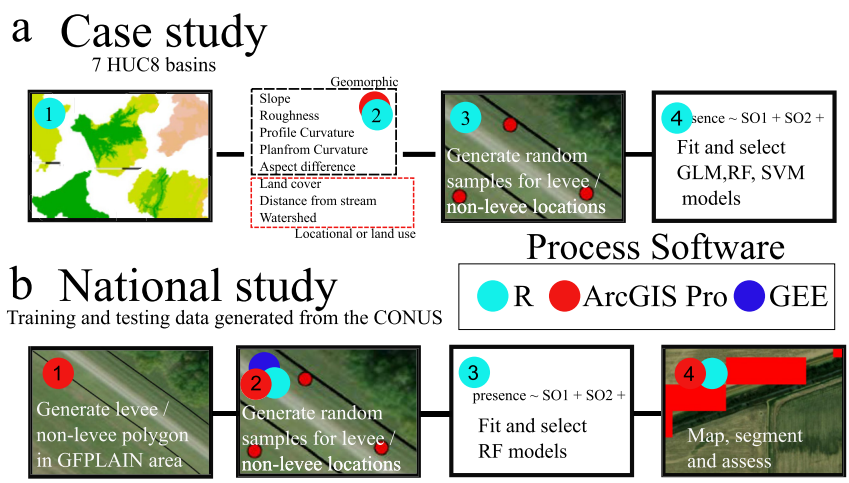

Figure 2. Workflow for the case and national studies. (a) In the case study we (1) generated 7 individual multi-layered rasters for each HUC8 basin hydrogeomorphic floodplain where each layer is a variable (2). (3) We combined random samples taken from levee (the NLD) and non-levee locations and (4) fit different GLM, RF, and SVM models as discussed above. We then assessed results. (b) In the national study, we (1) generated polygons around the NLD levees and all other areas in the hydrogeomorphic floodplain in each HUC2 basin to assist random sampling in ArcGIS Pro. (2) We generated random samples from the CONUS for two different sized data sets (n = 30,600 and 3,060,000) using GEE to generate the geomorphic variable values and ArcGIS Pro to generate the NLCD and distance from stream order 1–6 variables. (3) We fit different RF models to the data and applied the highest performing model to the contiguous U.S. hydrogeomorphic floodplain (4). We generated data from each HUC2 basin in ArcGIS Pro, imported and fit the data in R, and exported the model results back into ArcGIS Pro where we mapped, segmented, and then analyzed the results.

2016 NLCD, the six distance-from-stream-order variables, and the HUC8 basin (Table 2). We also tested models with each of the variables removed from the full model and some other combinations.

We analyzed the impacts of sample size, ratio of levee to non levee (presence to absence) data, and model variables on GLM, RF, and SVM performance through separate comparisons of the performance of models generated from a large number of data sets in which only those (e.g., sample size) quantities varied. We used 113 independent data sets varying from 110 to 13,900 sampled locations per basin with model 1. Based on those results, in

Table 3
Model Names and Variables in the Study

Model	Total variables	Variables
1	13	Geomorphic variables (slope, profile curvature, planform curvature, relative elevation, aspect difference), NLCD, basin, distance from stream order 1–6
2	12	Model 1 without slope
3	12	Model 1 without profile curvature
4	12	Model 1 without planform curvature
5	12	Model 1 without relative elevation
6	12	Model 1 without NLCD
7	12	Model 1 without aspect difference
8	12	Model 1 without basin
9	7	Model 1 without distance from stream order variables
10	6	Distance from stream order variables only
11	11	Model 1 without distance from stream order variables and aspect difference
12	8	NLCD, basin, distance from stream order 1–6

which the RF model outperformed SVM and GLM models at every sample size, we narrowed subsequent case study analyses to RF models only though we revisited all three model types at the CONUS level. The impacts of varying ratios of levee to non levee data to RF performance were analyzed using 93 independent data sets with absence to presence ratios ranging from 0.04 to 23.6. We attempted to hold sample size constant but experienced a range of 812 to 856 total sampled locations for each dataset due to varying "NA" value generation resulting from imperfect replacement of NA values sampled from masks. Absences were sampled everywhere in the GFPLAIN floodplain, excluding a 10 m buffer from the NLD centerlines. RF models with absence to presence ratios of 0.7 had the highest performance. Accordingly, we used this ratio for subsequent analyses although we tested ratios of 0.7 and 1.0 at the CONUS scale. We analyzed RF model performance with different variables using 50 independent data sets each consisting of data from 1,000 locations per basin with a 0.7 absence to presence ratio. Training data consisted of 70% of each data set with 30% used for validation. The results of these processes guided the CONUS study.

We conducted a leave-one-out cross-validation (Stone, 1974) with 1,171 NLD levees in the lower Mississippi Basin to understand how a parsimonious RF model (model 12) using the NLD and distance-to-stream-order variables behaved with undocumented levees that were not in the NLD. Given the close proximity of artificial levees (e.g., 74% of NLD levees in the LMR basin are within 5 km of each other) and the size of data sets (n > 3,000,000) used in the CONUS study, this cross validation best approximates a model's ability to detect undocumented levees. Using a 170,000-location sample with a 0.7 absence to presence ratio, we added a "levee ID" to each levee and wrote R code to iterate through each levee by first removing that levee from the data set, generating a model from 70% of the remaining data, applying the model to the withheld levee data, and recording the model accuracy. This process was conducted for each of the 1,171 artificial levees. We also recorded the shortest distance between each levee to understand how proximity impacted model performance.

2.5.2. National Modeling

In the national study, we tested different model types (GLM, RF, and SVM) and model variables using two different sized datasets generated from the CONUS floodplain for training and validation. We selected the most accurate model based on multiple accuracy metrics, a RF model trained on 2,142,000 sampled locations with land cover, HUC2 basin, and six distance from stream order 1–6 variables and applied it to the entire CONUS floodplain using ArcGIS Pro and R.

We used ArcGIS Pro to generate two polygons in each of the 18 HUC2 basins: a "levee" polygon within 10 m of the NLD, and a "non-levee" polygon for other areas in the hydrogeomorphic floodplain. Based on analyses in the case study, we randomly generated two differently sized data sets both with 0.7 absence to presence ratios: (a) 700 non-levee and 1000 levee locations in each HUC2 basin (total $n \sim 30{,}600$) and a much larger sample with (b) 70,000 non-levee and 100,000 levee locations in each HUC2 basin (total $n = 3{,}060{,}000$) for all 13 variables required for model 1. We used 70% of each data set for model training ($n = 21{,}420$ sampled locations for the small data set and 2,142,000 sampled locations for the large data set). We used GEE cloud computing and the Terrain Analysis in Google Earth Engine (TAGEE) script (Safanelli et al., 2020) to calculate the five geomorphic variables (Table 2). The land cover variable and the distance-from-stream-order variables were generated in ArcGIS Pro. These data were exported into R and fit to different models (Table 4).

We selected the best performing model based on a full suite of accuracy measurements (Table 4). We then generated a 30 m resolution raster from the calibrated CONUS 100-year hydrogeomorphic floodplain containing ~880,000,000 pixels. Working by HUC2 basin, we generated the dataset for each location in ArcGIS Pro, exported the data into R, applied the model, and exported the predicted values (in the case of RF models "1" or "0") back into ArcGIS Pro, where the predicted values were mapped and segmented. Segments were analyzed to determine whether they were already represented in the NLD, which stream order they were meant to "protect" against, and to estimate their length using segment attributes.

3. Results

3.1. Case Studies

RF models demonstrated the best predictive performance for identifying artificial levees at every sample size, followed by SVM models (Figure 3a). The GLMs demonstrated the worst performance of all three models.

Table 4
Model Performance in the National Study

Model[a]	ML[b]	Size[c]	Ratio[d]	k[e]	TP Rate[f]	TN Rate[g]	FP Rate[h]	FN Rate[i]
1	RF	1700	0.7	0.69	0.87	0.82	0.18	0.13
1	RF	1700	1	0.69	0.82	0.88	0.12	0.18
1	GLM	1700	1	0.47	0.68	0.80	0.20	0.32
1	SVM	1700	1	0.55	0.83	0.74	0.26	0.17
12	RF	1700	0.7	0.65	0.87	0.77	0.23	0.13
1	RF	170,000	0.7	0.85	0.95	0.90	0.10	0.05
12	RF	170,000	0.7	0.94	0.99	0.95	0.05	0.01
12 + relative elevation	RF	170,000	0.7	0.89	0.97	0.92	0.08	0.03
12 + profile curvature	RF	170,000	0.7	0.90	0.98	0.93	0.07	0.02
12 + aspect difference	RF	170,000	0.7	0.87	0.97	0.91	0.09	0.03
12 + slope	RF	170,000	0.7	0.89	0.97	0.92	0.08	0.03

[a]"Model" corresponds to the variables listed in Table 3. [b]The "ML" column denotes the machine learning or statistical model used. [c]"Size" denotes the total sample size taken from each HUC2 basin for both model training and testing. [d]"Ratio" denotes the ratio of absence to presence in the sample. [e]The result denotes the Cohen's kappa of the model on the testing sample, where we used a 70/30 random split for training and validation in all models. [f]True positive rate. [g]True negative rate. [h]False positive rate. [i]False negative rate.

Because the RF models demonstrated the best performance, all our subsequent results in the case study focus on RF model outputs. An absence-to-presence ratio of 0.7 resulted in the best RF model performance, but data with ratios between 0.45 and 1.24 performed well (Figure 3b). Furthermore, RF performance continued to improve with larger sample sizes from each basin, exceeding a kappa of 0.8 for sample sizes larger than 1,000.

Different RF models, each with 100 trees of three variables sampled at each node, were applied to 50 different random samples from 1,000 sampled locations and a 0.7 absence/presence ratio (Figure 4). Model 1, with all variables, only slightly outperformed models with one less variable with kappas in the 0.75–0.8 range. A model without any geomorphic variables (model 12) performed almost as well as the full model.

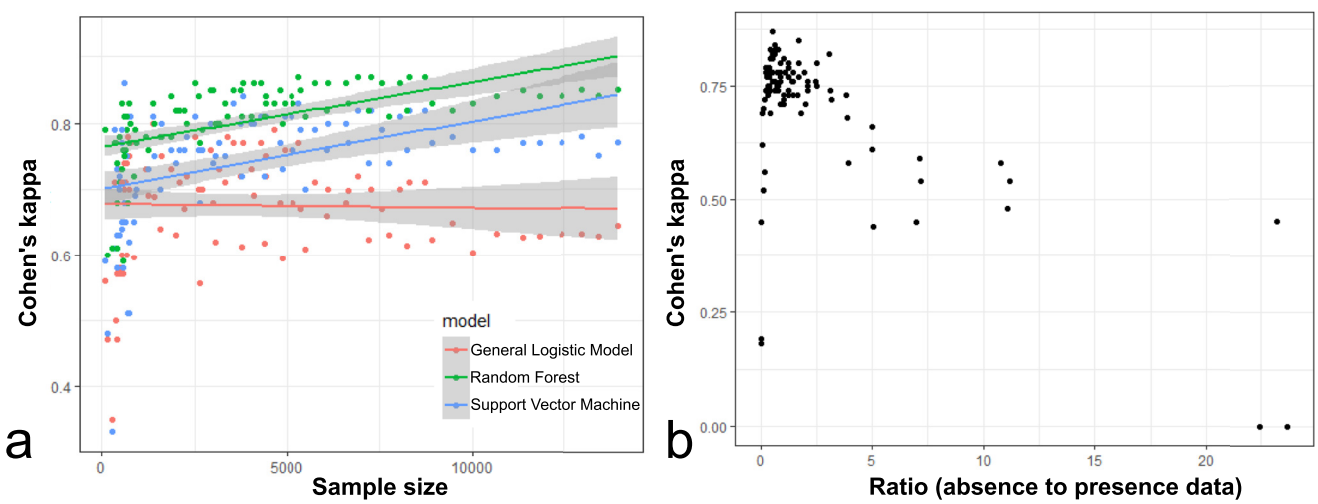

Figure 3. Model performance using Cohen's kappa in the case study. (a) GLM, RF, and SVM model performance with sample size varying from 110 to 13,900 sampled locations total in the seven basins for 113 independent samples. The gray envelopes are 95% confidence intervals for logistic models, depicted by a solid line, fit to the data (b) Performance of 93 independent RF models by varying absence/presence ratio of sampled locations while controlling for sample size ($n \sim 832$ sampled locations).

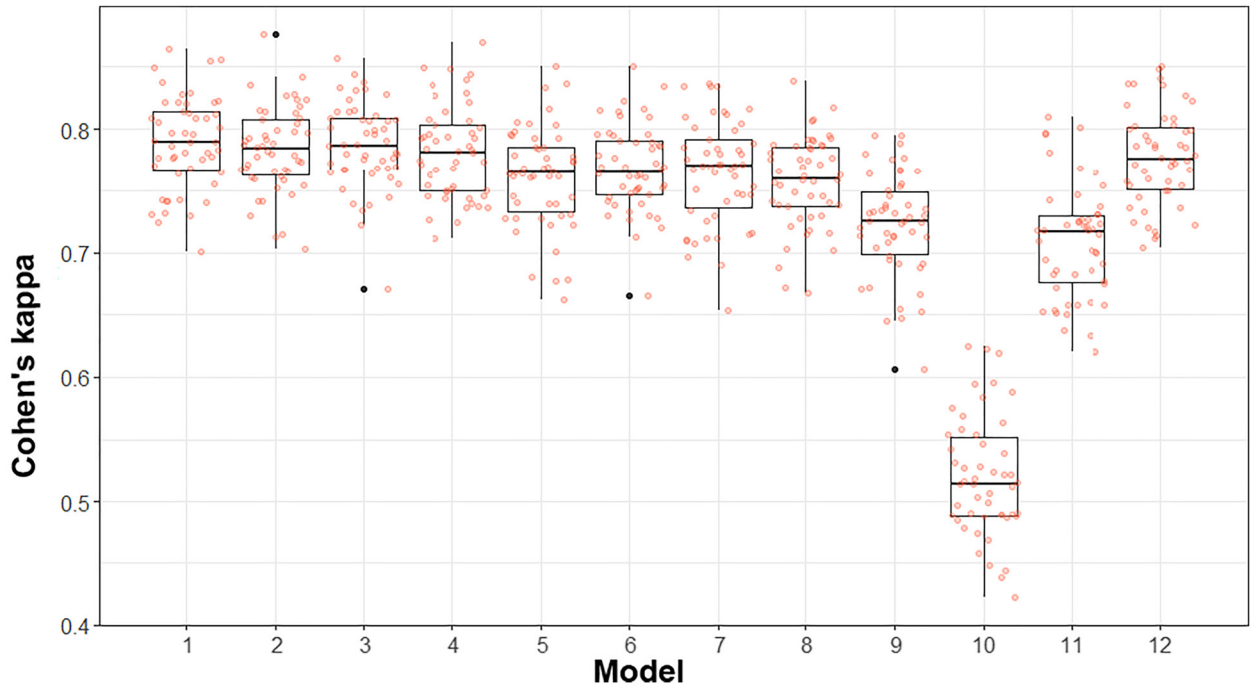

Figure 4. RF model performance by Cohen's kappa and variables for 50 data sets, each with a 0.7 ratio of absence to presence data and ~1,000 sampled locations. Boxplots are plotted along with individual model values. The model number on the *x*-axis corresponds to models listed in Table 2.

3.2. Leave-One-Out Cross-Validation

An RF model using model 12 detected 61% of levees when they were left out of the training dataset. Detected levees were longer than undetected levees such that sum of the length of detected levees (7,473 km) represented 94% of total levee length (7,910 km; Figure 5a). Levees were close together, with 74% of levees within 5 km of each other and 94% of levees within 25 km (Figure 5b).

3.3. National Study

We tested different variables, model types, sample sizes and absence/presence ratios using models trained at a national scale (Table 4). As in the case study, RF models outperformed SVMs and GLMs with RF model performance increasing with sample size. We were initially surprised to discover that model 12, without any of the five geomorphic variables, outperformed the model 1 by 0.1 kappa and other performance measures (Table 4, Figure S2 in Supporting Information S1). Further, we did not initially test model 12 in the case study, and only added it after discovering it here.

The performance of model 12 corresponds to a 97% accuracy, meaning that 97% of the known non-levee and levee locations in the ~918,000 CONUS validation sample were predicted correctly, although the results of this study indicate that many false-positives may represent undocumented artificial levees (what we call potential levees). However, the model performance varied spatially, ranging from 0.813 kappa in the California HUC2 basin to 0.999 kappa in the Upper Colorado HUC2 basin, correlating to the inverse of the NLD levee total length in each HUC2 basin.

Potential artificial levees, those areas identified that may be artificial levees and are not identified by the NLD, are listed in Table 5 (which also lists NLD levees separately). Potential levees were concentrated in the upper and lower Mississippi and the Missouri basins (basins 7, 8, 10 in Table 5 and Figure 6). Potential levees were also concentrated along streams of order 2 to 6, constituting 75% of total levee length (Figure 7). There were 146,404 potential levees identified constituting a total length of 182,213 km (Table 5). Normalized artificial levee length (the total length of NLD or potential levees associated with stream order X divided by the combined length of streams with stream order X) gives a sense of how streams of a particular order are impacted by artificial levees.

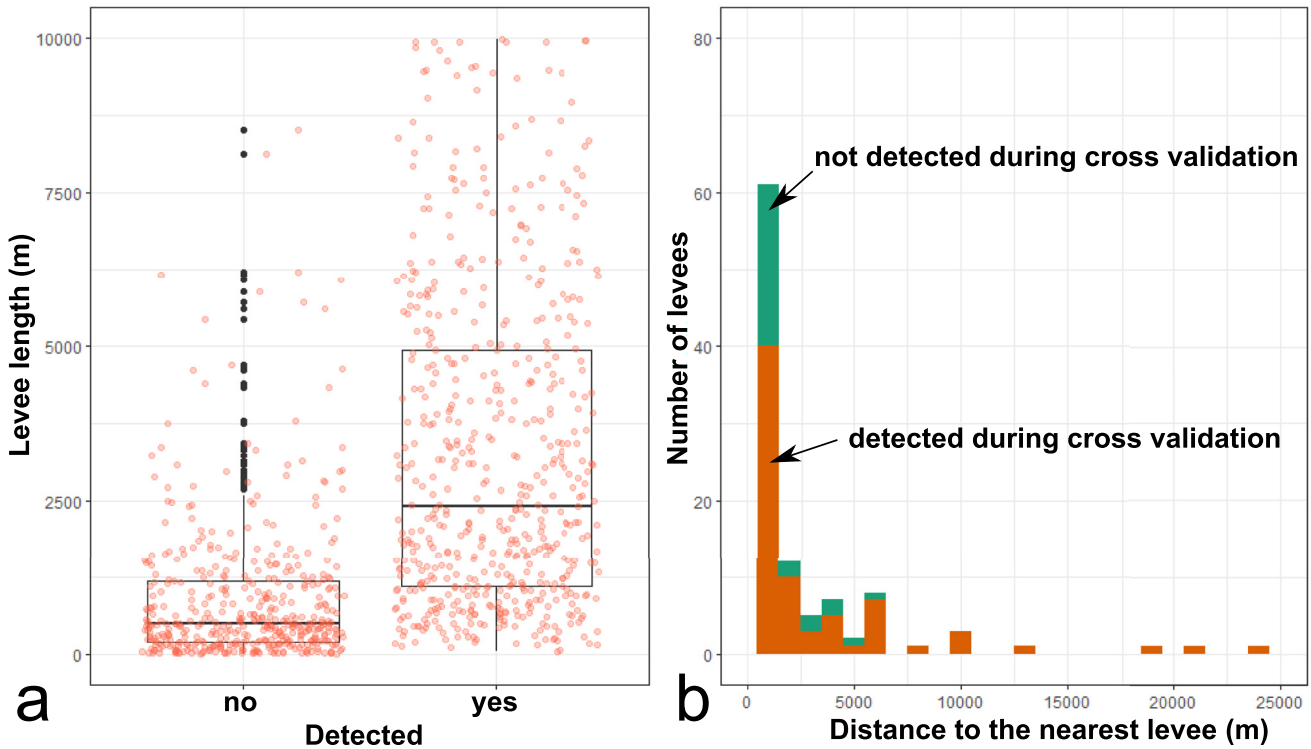

Figure 5. Results from the leave-one-out cross-validation. (a) Longer levees were detected more often than shorter levees so that detected levees represent 94% of total levee length. (b) Levees are close together, with 74% of levees within 5 km of each other and 94% within 25 km.

Artificial levees provide greater "protection" along streams of greater order, with normalized length approaching 0.20 for stream order 10 (Figure 7). Potential levees and those documented in the NLD represent coverage of 2% of the total length of streams in the contiguous United States.

To illustrate a few locations where we identified levees not present in the NLD and to discuss how the model works, we highlight three potential levees and the raw model results (prior to segmentation) that we were able to ascertain are definitely levees (Figure 8). Brookville, Indiana sits at the confluence of the East Fork Whitewater River (stream order 7) and the Whitewater River (stream order 8), which is a tributary to the Ohio River (Figure 8a). This artificial levee was built along the East Fork Whitewater River and is more than 850 m long, 1–2 m tall, and has been the subject of a riverwalk project (Norwood, 2020). This illustrates how undocumented artificial levees influence hydrogeomorphic floodplains, as the 100-year floodplain and detection are immediately next to the levee. Elaine, Arkansas is situated several kilometers from the Mississippi River next to an old meander scar and is "protected" from Mississippi River floods by kilometers of massive artificial levees. One 400 m section of artificial levee is not documented in the NLD (Figure 8b). Arcata, California is a tidally influenced community situated next to Arcata Bay and the Pacific Ocean in northern California. The area was subject to salt marsh reclamation for pasturage during the last 150 years (Murray & Wunner, 1980). This 200 m length of artificial levee, visible from U.S. Route 101, was likely built to reclaim floodplain along Jacoby Creek for pasturage (Figure 8c). Detected in the raw model results, this large segment was not included in the final results because the segment overlapped a NLD levee to the north. These results indicate the value of buffering the hydrogeomorphic floodplain by 100 m and considering how to separate model results from the NLD.

4. Discussion

4.1. Location and Prevalence of Artificial Levees

Our analysis indicates that the NLD may only show 20.4% of the artificial levees in the CONUS. Over 62% of potential levees are concentrated in the upper and lower Mississippi basins and the Missouri basins, with potential levee length exceeding documented levee length by factors of seven, five, and nine, respectively

Table 5
NLD and Potential Levee Lengths (km) by HUC2 Basin

HUC2 basin	NLD (km)	Potential levees (km)
New England (1)	89	25
Mid-Atlantic (2)	617	2,220
South Atlantic-Gulf (3)	2,410	5,921
Great Lakes (4)	41	277
Ohio (5)	1,148	12,216
Tennessee (6)	45	40
Upper Mississippi (7)	4,804	35,374
Lower Mississippi (8)	7,912	38,657
Souris-Red-Rainy (9)	466	459
Missouri (10)	4,438	39,221
Arkansas-White-Red (11)	2,939	16,073
Texas-Gulf (12)	2,403	9,230
Rio Grande (13)	1,074	965
Upper Colorado (14)	154	9
Lower Colorado (15)	1,582	1,471
Great Basin (16)	133	392
Pacific Northwest (17)	2,082	10,747
California (18)	14,306	8,916
Total Length (km)	46,643	182,213

Note. Potential levees were identified by this study and are not part of the NLD.

(Figure 6, Table 5). Potential levee length in the Ohio basin exceeds NLD levees by a factor of 11. These details, combined with the concentration of potential levees along smaller stream orders (Figure 7), seems to reflect the long history of artificial levee construction in the lower Mississippi River basin and early interventions along the Mississippi and Ohio Rivers, such as the 1824 Rivers and Harbors Bill (Wohl, 2005). We suspect that NLD and potential levees represent a continuum in which NLD levees represent one pole with larger levees more recently built with state or federal funds, and designed to "protect" against streams of a larger order. The other pole is represented by potential levees which are smaller, built longer ago and by landowners, and situated next to smaller streams. We see this dynamic illustrated in the California basin, where potential levee length is half that of NLD levee length (Table 5), reflecting the later period (i.e., early twentieth century) of artificial levee construction in that basin (ASCE, 2017) and better documentation that comes with more recent construction.

4.2. Spatial and Geographic Implications of Findings

Characterizing the shape of Earth's surface is considered the primary method for quantitative land-surface analysis (Sofia, 2020). Our understanding of landscapes arises from human cognition of spatial patterns in the field (Roering et al., 2013), so it is reasonable that geomorphic variables are nearly the exclusive set of variables used in recent efforts to detect artificial levees and other earthworks. Nonetheless, topographic patterns are only one way to parameterize geomorphic features. A quick review of efforts to model other physical phenomena with less pronounced topographic profiles, such as wetlands, indicates a willingness of researchers to explore other ancillary variables such as "distance to stream" (e.g., Ansari & Golabi, 2019; Berhane et al., 2018; Golden et al., 2016). Unlike other geomorphic features of mostly natural origin, artificial levees are constructed solely by humans to protect infrastructure from river floods. They are anthropogenic features that are intimately tied to human land cover and hydrologic features. Consequently, the spatial patterns humans have created in building levees are real and useful for modeling.

We are interested in the causes of the difference in accuracy between models that use land cover and stream order variables (e.g., model 12) and those models using geomorphic variables (e.g., model 1). Detecting artificial levees presents significant technical challenges due to their small size, geographic ubiquity, and varied morphology (Steinfeld & Kingsford, 2013). Artificial levees can be massive structures or features nearly invisible to both the eye and topographically based analyses (Figure S3 in Supporting Information S1), with the height of some artificial levees less than the vertical error of topographic datasets (e.g., the mean relative vertical accuracy of the NED is 0.81 m with the accuracy of 95% of locations within 2.93 m (Gesch et al., 2014)). Furthermore, the resampling process of digital elevation models tends to smooth topographic crests (such as those of levees) making the features even more topographically stealthy or even invisible (Wing et al., 2019). Consequently, it is not surprising that spatial and land cover patterns seem to be more useful than geomorphic patterns in a national study given the diverse geomorphic signatures of both documented artificial levees (such as those in Figure S3 in Supporting Information S1, which can be used as training data) and undocumented levees. Other researchers working at the national scale or larger (e.g., Grill et al., 2019) recognized the strong correlation of lateral discontinuity structures with human development and employed nightlight intensity as a proxy for lateral discontinuity in the absence of global records of artificial levees and other structures.

Recent investigations have raised concerns over validation strategies for large scale modeling studies where the employment of spatially autocorrelated training and validation data leads to inflated estimates of model accuracy (e.g., Ploton et al., 2020). The issue is covered extensively, especially in the ecological literature (see Roberts et al., 2017), but we consider it appropriate to discuss the suitability of the validation techniques employed here. Karasiak et al. (2021) explains how the winner of a land cover classification contest was able to employ geographic

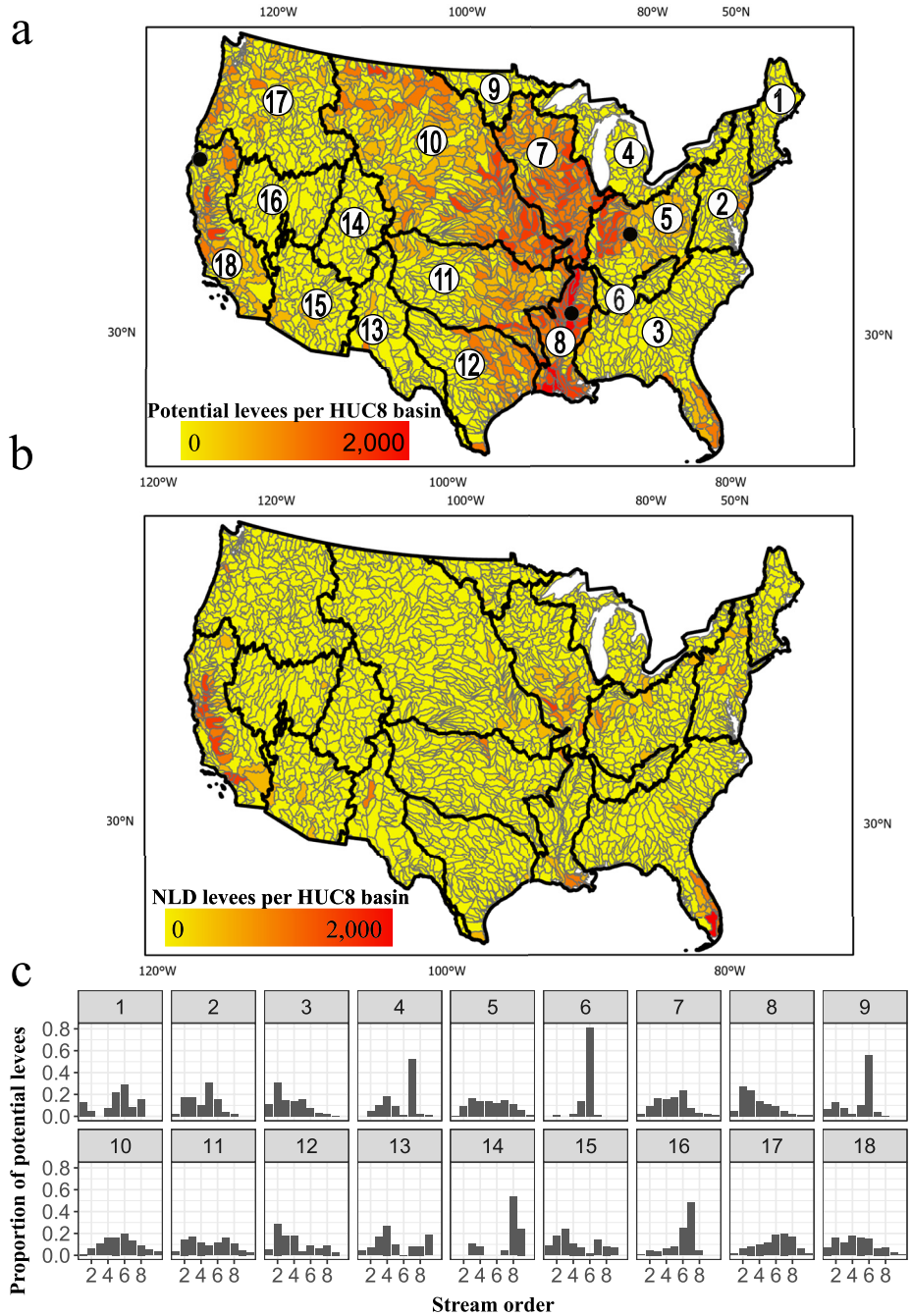

Figure 6. A spatial and stream order representation of potential and NLD artificial levees by HUC8 and HUC2 basin. (a) The number of potential levees per HUC8 basin. HUC2 basin boundaries, in bold, are denoted by number. Three black dots indicate potential levees examined in Figure 8. (b) The number of NLD levees per HUC8 basin with HUC2 basin boundaries in bold. (c) The proportion of potential artificial levee length along each stream order in 18 HUC2 basins.

pixel location only due to the ostensible spatial autocorrelation of the training and validation data sets. Unlike that example, and as previously discussed, we consider the spatial patterns expressed by the distance from stream order and land cover variables to be real patterns created by humans because land cover and stream flow were primary factors in the decision process that led to artificial levee construction. In addition, our method of mapping model 12 over the GFPLAIN floodplain is considered interpolation, not extrapolation, because we are applying the model in the same domain (i.e., the same geographic extent and variable domain) as that from which the training data are generated. Validation error of random samples is considered accurate in models with applications

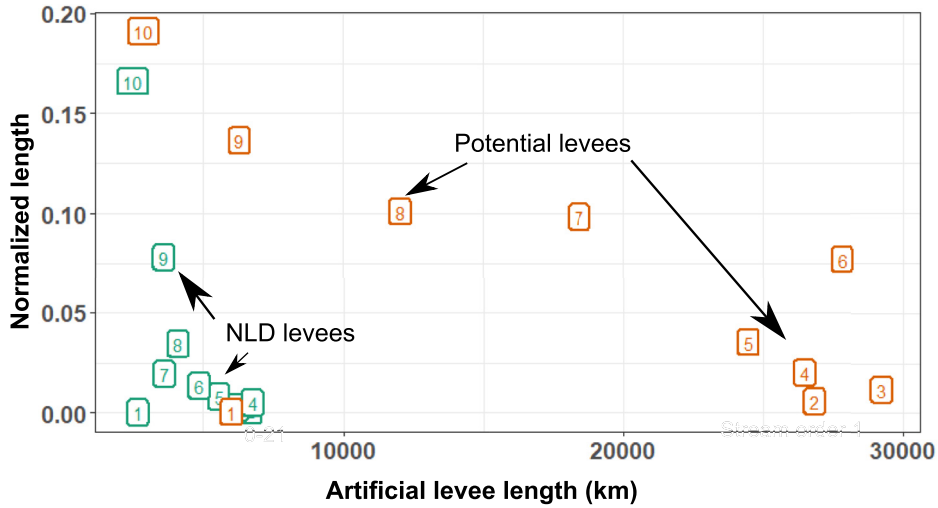

Figure 7. Potential and NLD artificial levee stream orders by normalized length and the sum of levee length for that order. NLD and potential levees are annotated by green and orange colors, respectively.

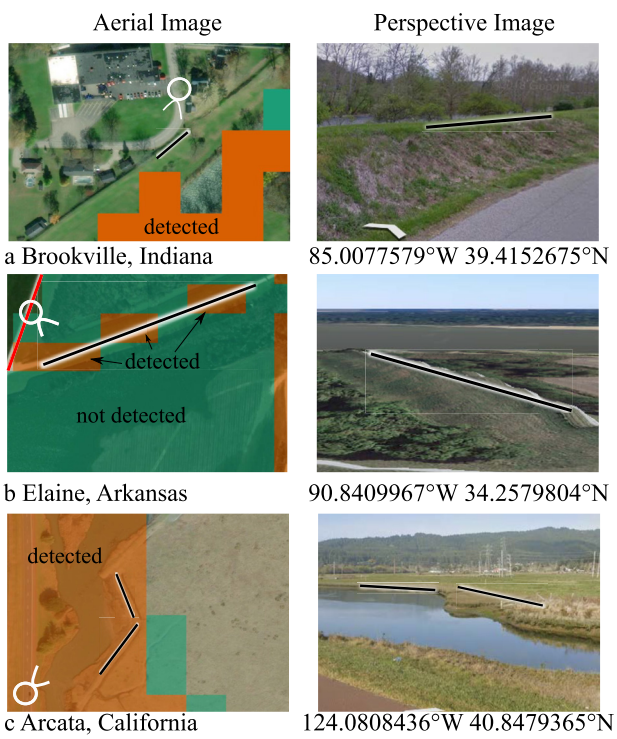

Figure 8. Aerial and perspective images of potential artificial levees discovered during this process. The white circle and arrow on the aerial image indicates the perspective location for the perspective image. Black lines correlate locations between the two images. Raw model results, prior to segmentation, are plotted in orange (detected) and green (not detected). (a) Brookville, Indiana levee is visible as a long linear feature in aerial imagery, on Google Street view, and written about in an online news article (Norwood, 2020). (b) Elaine, Arkansas levee on the Mississippi River is connected to a NLD levee (indicated by a red line) but remains undocumented. (c) Arcata, California levee along the Gannon Slough and U.S. Route 101 is likely related to salt marsh reclamation for pasturage.

in similar geographic and variable domains (Roberts et al., 2017). Unlike the problematic models discussed by Ploton et al. (2020), our training and validation samples ($n \sim 3{,}060{,}000$) are drawn from the same geographic and variable space as the model application area (the full 100-year GFPLAIN floodplain). We are not applying the model in a different geographic area. The detection of unknown levees representing 94% of total levee length in the leave-one-out cross-validation substantiates these claims. Initially, our use of the 2-digit HUC code, instead of the 8-digit code, as a variable was driven by the RF package's (Liaw & Wiener, 2002) limit of 32 levels for a factor variable. Our experience indicates that the use of the 18 levels of the 2-digit HUC code strikes the right balance between harnessing RF's ability to model large samples without overfitting (Breiman, 2001) and the need to apply spatial patterns locally. After even a casual perusal of the USACE NLD website (https://levees.sec.usace.army.mil), it is reasonable to suspect that different areas of the United States might exhibit different spatial patterns of artificial levees and thus employing the natural basin configuration at the regional scale is appropriate. An example of this is the artificial levee density difference between basins such as the lower Mississippi Basin and California Basin (with a relatively high density of artificial levees) and the New England basin (with a relatively low density). Another consideration for modeling is that the 2-digit HUC basin level is the most specific HUC code designation so that all basins contain known levees. CONUS models using HUC4 basins or larger HUC code designators would have basins without training data.

4.3. Future Directions

Several areas appear promising for future research. First, artificial levee detection from a truly object-based approach could allow for the introduction of several object-based variables such as levee length and volume, mimicking the methods that experienced engineers or geomorphologists might employ to judge whether a structure is an artificial levee. Second, including additional variables, such as stream order 7–10 variables and spectral properties, could improve model accuracy. To test additional variables and the idea of using a larger sample size, we added a larger training sample size and a distance from

Table 6
Model Performance in the Lower Mississippi HUC2 Basin With a Larger Training Sample Size and an Additional Variable

Model	*n*	Kappa	TP Rate[a]	TN Rate[b]	FP Rate[c]	FN Rate[d]
12	170,000	0.846	0.94	0.91	0.09	0.06
12	1,700,000	0.958	0.97	0.99	0.01	0.03
12 + SO10	1,700,000	0.974	0.98	0.99	0.01	0.02

[a]True positive rate. [b]True negative rate. [c]False positive rate. [d]False negative rate.

stream order 10 variable in the lower Mississippi Valley (Table 6, Figure S4 in Supporting Information S1). Both ideas improved model accuracy, with increased sample size resulting in the greatest gains.

Third, we suggest coupling a geomorphic post-processing technique using high resolution topographic data with this model to determine whether the areas identified as potential-levees are shaped like levees. Although we found that including geomorphic variables reduced predictive power at the national scale, assessing how "levee like" potential levees are geomorphically in a subsequent step could help prioritize ground-truthing efforts. Fourth, the incorporation of field validation of potential levees similar to the validation of barriers conducted by Jones et al. (2019) would improve model certainty. Actual ground-truthing of potential levees identified here could complement the methods currently in use to document artificial levees for the NLD, resulting in a more complete and certain database. Fifth, revisiting this analysis in a decade could benefit from expanded computing ability and more open source options for object based classification, a more thorough NLD as training data, and a more accurate NHD Plus HR. Sixth, the causes of the spatial patterns of potential and NLD levees, illustrated in Figure 6, deserve more detailed exploration.

5. Conclusions

Our exploration of different variables and models to detect artificial levees led to a random forest model with land cover and stream order variables. Applying this model in a 100-year geomorphic floodplain in the contiguous US indicated the potential for 182,000 km of artificial levees that are not included in the national levee database, suggesting that the database contains 20.4% of artificial levee length in the continental US. These levees missing from the national database were concentrated in the lower and upper Mississippi and Missouri basins and mostly along streams of order 2 through 6. When normalized for total stream length, larger stream orders were more impacted than smaller streams, with more than a third of stream order 10 streams impacted by NLD or potential levees. Ideas for future directions include methods that could further improve model performance and validation.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Potential levee location, estimated length, and HUC2 basin are available as supporting data (Knox et al., 2022): https://doi.org/10.4211/hs.729c0aea00bb48d6b6814c147e4318c4.

Acknowledgments
The authors thank Tracie-Lynn Nadeau, Shane Cline, Doug Thompson, and Francis Magilligan for suggesting basins to consider for the case study. Kevin McNinch's assistance in providing a digital copy of the NHD plus was critical. The authors thank Oliver Wing for providing MATLAB code to generate geomorphic variables. Antonio Annis's help in running GFPLAIN in ArcGIS Pro was extremely helpful. Ann Hess's statistical advice was invaluable. The manuscript was improved by comments from three anonymous reviewers.

References

American Society of Civil Engineers (2017). *2017 Infrastructure Report card: Levees, Reston, Va*. Retrieved From https://www.infrastructurereportcard.org/cat-item/levees/

Annis, A., Karpack, M., Morrison, R. R., & Nardi, F. (2022). On the influence of river basin morphology and climate on hydrogeomorphic floodplain delineations. *Advances in Water Resources*, *159*, 104078. https://doi.org/10.1016/j.advwatres.2021.104078

Annis, A., Nardi, F., Morrison, R. R., & Castelli, F. (2019). Investigating hydrogeomorphic floodplain mapping performance with varying DTM resolution and stream order. *Hydrological Sciences Journal*, *64*(5), 525–538. https://doi.org/10.1080/02626667.2019.1591623

Ansari, A., & Golabi, M. H. (2019). Prediction of spatial land use changes based on LCM in a GIS environment for Desert Wetlands–A case study: Meighan Wetland, Iran. *International soil and water conservation research*, *7*(1), 64–70. https://doi.org/10.1016/j.iswcr.2018.10.001

Barry, J. M. (2007). *Rising tide: The Great Mississippi Flood of 1927 and how it changed America* (p. 528). Simon and Schuster.

Berhane, T. M., Lane, C. R., Wu, Q., Autrey, B. C., Anenkhonov, O. A., Chepinoga, V. V., & Liu, H. (2018). Decision-tree, rule-based, and random forest classification of high-resolution multispectral imagery for wetland mapping and inventory. *Remote Sensing*, *10*(4), 580. https://doi.org/10.3390/rs10040580

Blanton, P., & Marcus, W. A. (2009). Railroads, roads and lateral disconnection in the river landscapes of the continental United States. *Geomorphology*, *112*, 212–227. https://doi.org/10.1016/j.geomorph.2009.06.008

Blaschak, T. S. (2012). *Examining geomorphic effects of flow diversions on low-gradient mountain streams in the Routt National Forest*. Doctoral dissertation, Colorado State University.

Blessing, R., Sebastian, A., & Brody, S. D. (2017). Flood risk delineation in the United States: How much loss are we capturing? *Natural Hazards Review*, *18*(3), 04017002. https://doi.org/10.1061/(asce)nh.1527-6996.0000242

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324
Briggs, M. A., Lautz, L. K., Hare, D. K., & González-Pinzón, R. (2013). Relating hyporheic fluxes, residence times, and redox-sensitive biogeochemical processes upstream of beaver dams. *Freshwater Science*, *32*(2), 622–641. https://doi.org/10.1899/12-110.1
Brooks, G. R. (2003). Holocene lateral channel migration and incision of the Red River, Manitoba, Canada. *Geomorphology*, *54*(3–4), 197–215. https://doi.org/10.1016/s0169-555x(02)00356-2
Brown, A. G., Tooth, S., Bullard, J. E., Thomas, D. S., Chiverrell, R. C., Plater, A. J., et al. (2017). The geomorphology of the Anthropocene: Emergence, status and implications. *Earth Surface Processes and Landforms*, *42*(1), 71–90. https://doi.org/10.1002/esp.3943
Bunn, S. E., Abal, E. G., Smith, M. J., Choy, S. C., Fellows, C. S., Harch, B. D., et al. (2010). Integration of science and monitoring of river ecosystem health to guide investments in catchment protection and rehabilitation. *Freshwater Biology*, *55*, 223–240. https://doi.org/10.1111/j.1365-2427.2009.02375.x
Buto, S. G., & Anderson, R. D. (2020). *NHDPlus High Resolution (NHDPlus HR)—A hydrography framework for the Nation (No. 2020–3033)*. US Geological Survey.
Cannon, C. M. (2015). *Landforms along the Lower Columbia River and the influence of humans*. Doctoral dissertation, Portland State University.
Caskey, S. T. (2013). *Downstream effects of diversion dams on riparian vegetation communities in the Routt National Forest*. Colorado State University.
Caskey, S. T., Blaschak, T. S., Wohl, E., Schnackenberg, E., Merritt, D. M., & Dwire, K. A. (2015). Downstream effects of stream flow diversion on channel characteristics and riparian vegetation in the Colorado Rocky Mountains, USA. *Earth Surface Processes and Landforms*, *40*(5), 586–598. https://doi.org/10.1002/esp.3651
Castro, J. M., & Thorne, C. R. (2019). The stream evolution triangle: Integrating geology, hydrology, and biology. *River Research and Applications*, *35*(4), 315–326. https://doi.org/10.1002/rra.3421
Cho, E., Jacobs, J. M., Jia, X., & Kraatz, S. (2019). Identifying subsurface drainage using satellite Big Data and machine learning via Google Earth Engine. *Water Resources Research*, *55*(10), 8028–8045. https://doi.org/10.1029/2019wr024892
Choi, C., Kim, J., Han, H., Han, D., & Kim, H. S. (2020). Development of water level prediction models using machine learning in wetlands: A case study of Upo wetland in South Korea. *Water*, *12*(1), 93.
Covino, T. (2017). Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. *Geomorphology*, *277*, 133–144. https://doi.org/10.1016/j.geomorph.2016.09.030
Criss, R. E., & Shock, E. L. (2001). Flood enhancement through flood control. *Geology*, *29*(10), 875–878. https://doi.org/10.1130/0091-7613(2001)029<0875:fetfc>2.0.co;2
Czech, W., Radecki-Pawlik, A., Wyżga, B., & Hajdukiewicz, H. (2016). Modelling the flooding capacity of a polish Carpathian river: A comparison of constrained and free channel conditions. *Geomorphology*, *272*, 32–42. https://doi.org/10.1016/j.geomorph.2015.09.025
Czuba, C., Williams, B., Westman, J., & LeClaire, K. (2015). *An assessment of two methods for identifying undocumented levees using remotely sensed data*. US Department of the Interior, US Geological Survey.
Deines, J. M., Kendall, A. D., & Hyndman, D. W. (2017). Annual irrigation dynamics in the US Northern High Plains derived from Landsat satellite data. *Geophysical Research Letters*, *44*(18), 9350–9360. https://doi.org/10.1002/2017gl074071
Esri Inc. (2020). *ArcGIS Pro*. ESRI Inc. Retrieved from https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview
Fitzgerald, R. W., & Lees, B. G. (1994). Assessing the classification accuracy of multisource remote sensing data. *Remote Sensing of Environment*, *47*(3), 362–368. https://doi.org/10.1016/0034-4257(94)90103-1
Foody, G. M. (2020). Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment*, *239*, 111630. https://doi.org/10.1016/j.rse.2019.111630
Frings, R. M., Berbee, B. M., Erkens, G., Kleinhans, M. G., & Gouw, M. J. (2009). Human-induced changes in bed shear stress and bed grain size in the River Waal (The Netherlands) during the past 900 years. *Earth Surface Processes and Landforms*, *34*(4), 503–514. https://doi.org/10.1002/esp.1746
Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., & Tyler, D. (2002). The National Elevation Dataset. *Photogrammetric Engineering & Remote Sensing*, *68*(1), 5–32.
Gesch, D. B., Oimoen, M. J., & Evans, G. A. (2014). *Accuracy Assessment of the US Geological Survey National Elevation Dataset, and Comparison With Other Large-Area Elevation Datasets: SRTM and ASTER* (Vol. 1008). US Department of the Interior, US Geological Survey.
Golden, H. E., Sander, H. A., Lane, C. R., Zhao, C., Price, K., D'Amico, E., & Christensen, J. R. (2016). Relative effects of geographically isolated wetlands on streamflow: A watershed-scale analysis. *Ecohydrology*, *9*(1), 21–38. https://doi.org/10.1002/eco.1608
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031
Graf, W. L. (1999). Dam nation: A geographic census of American dams and their large-scale hydrologic impacts. *Water Resources Research*, *35*(4), 1305–1311. https://doi.org/10.1029/1999WR900016
Graf, W. L. (2001). Damage control: Restoring the physical integrity of America's rivers. *Annals of the Association of American Geographers*, *91*(1), 1–27. https://doi.org/10.1111/0004-5608.00231
Grill, C., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019). Mapping the world's free-flowing rivers. *Nature*, *569*, 215–221. https://doi.org/10.1038/s41586-019-1111-9
Guilliams, D. (1998). The Big Ditch: The Wichita-Valley center flood control project. *Fairmount Folio: Journalism History*, *2*.
Harvey, J., & Gooseff, M. (2015). River corridor science: Hydrologic exchange and ecological consequences from bedforms to basins. *Water Resources Research*, *51*(9), 6893–6922. https://doi.org/10.1002/2015WR017617
Heine, R. A., & Pinter, N. (2012). Levee effects upon flood levels: An empirical assessment. *Hydrological Processes*, *26*(21), 3225–3240. https://doi.org/10.1002/hyp.8261
Hirschboeck, K. K. (1991). Climate and floods. *U. S. Geological Survey Water-Supply Paper*, *2375*, 67–88.
Horritt, M. S., & Bates, P. D. (2001). Effects of spatial resolution on a raster based model of flood flow. *Journal of Hydrology*, *253*(1–4), 239–249. https://doi.org/10.1016/s0022-1694(01)00490-5
Hudson, P. F., Middelkoop, H., & Stouthamer, E. (2008). Flood management along the Lower Mississippi and Rhine Rivers (The Netherlands) and the continuum of geomorphic adjustment. *Geomorphology*, *101*(1–2), 209–236. https://doi.org/10.1016/j.geomorph.2008.07.001
Jackson, J. K., Huryn, A. D., Strayer, D. L., Courtemanch, D. L., & Sweeney, B. W. (2005). Atlantic Coast Rivers of the northeastern United States. In *Rivers of North America* (pp. 20–71). Academic Press. https://doi.org/10.1016/b978-012088253-3/50005-5
Jin, S., Homer, C., Yang, L., Danielson, P., Dewitz, J., Li, C., et al. (2019). Overall methodology design for the United States national land cover database 2016 products. *Remote Sensing*, *11*(24), 2971. https://doi.org/10.3390/rs11242971
Jones, J., Börger, L., Tummers, J., Jones, P., Lucas, M., Kerr, J., et al. (2019). A comprehensive assessment of stream fragmentation in Great Britain. *Science of the Total Environment*, *673*, 756–762. https://doi.org/10.1016/j.scitotenv.2019.04.125

Junk, W. J., Bayley, P. B., & Sparks, R. E. (1989). The flood pulse concept in river-floodplain systems. In *Proceedings of the international large river symposium* (Vol. 106, pp. 110–127). Canadian Special Publications in Fisheries and Aquatic Sciences.
Karasiak, N., Dejoux, J. F., Monteil, C., & Sheeren, D. (2021). Spatial dependence between training and test sets: Another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*, 1–25. https://doi.org/10.1007/s10994-021-05972-1
Katibah, E. F. (1984). A brief history of riparian forests in the Central Valley of California. In *California riparian systems* (pp. 23–29). University of California Press. https://doi.org/10.1525/9780520322431-008
Knox, R., Morrison, R., & Wohl, E. (2022). *Potential levees*. HydroShare. https://doi.org/10.4211/hs.729c0aea00bb48d6b6814c147e4318c4
Knox, R. L., & Latrubesse, E. M. (2016). A geomorphic approach to the analysis of bedload and bed morphology of the Lower Mississippi River near the Old River Control Structure. *Geomorphology*, *268*, 35–47. https://doi.org/10.1016/j.geomorph.2016.05.034
Kondolf, G. M., Boulton, A. J., O'Daniel, S., Poole, G. C., Rahel, F. J., Stanley, E. H., et al. (2006). Process-based ecological river restoration: Visualizing three-dimensional connectivity and dynamic vectors to recover lost linkages. *Ecology and Society*, *11*(2). https://doi.org/10.5751/es-01747-110205
Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, *7*(1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003
Lehner, B., Liermann, C. R., Revenga, C., Vorosmarty, C., Fekete, B., Crouzet, P., et al. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, *9*, 494–502. https://doi.org/10.1890/100125
Leopold, L. B., & Maddock, T. (1953). *The Hydraulic Geometry of Stream Channels and Some Physiographic Implications* (Vol. 252). US Government Printing Office.
Liaw, A., & Wiener, M. (2002). *Classification and regression by random forest* (Vol. 2, pp. 18–22). R News.
Lindersson, S., Brandimarte, L., Mård, J., & Di Baldassarre, G. (2021). Global riverine flood risk–how do hydrogeomorphic floodplain maps compare to flood hazard maps? *Natural Hazards and Earth System Sciences Discussions*, 1–38.
Matthews, W. J., Vaughn, C. C., Gido, K. B., & Marsh-Matthews, E. D. I. E. (2005). *Southern Plains Rivers* (pp. 282–325). Rivers of North America. https://doi.org/10.1016/b978-012088253-3/50010-9
Murray, A., & Wunner, R. (1980). *A Study of the Jacoby Creek Watershed, Humboldt County, California*. Jacoby Creek Canyon Community.
Nardi, F., Annis, A., Di Baldassarre, G., Vivoni, E. R., & Grimaldi, S. (2019). GFPLAIN250m, a global high-resolution dataset of Earth's floodplains. *Scientific Data*, *6*(1), 1–6. https://doi.org/10.1038/sdata.2018.309
Nardi, F., Morrison, R. R., Annis, A., & Grantham, T. E. (2018). Hydrologic scaling for hydrogeomorphic floodplain mapping: Insights into human-induced floodplain disconnectivity. *River Research and Applications*, *34*(7), 675–685. https://doi.org/10.1002/rra.3296
Nardi, F., Vivoni, E. R., & Grimaldi, S. (2006). Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. *Water Resources Research*, *42*(9). https://doi.org/10.1029/2005WR004155
Norwood, J. (2020). *Riverwalk project dies in Brookville. Tri-County Weekend*. Retrieved from https://m.thecourierexpress.com/tri_county_sunday/news/local/riverwalk-project-dies-in-brookville/article_34a0f4ec-2da1-5b78-863e-26468c50b3a9.html
Palmer, M. A., Hondula, K. L., & Koch, B. J. (2014). Ecological restoration of streams and rivers: Shifting strategies and shifting goals. *Annual Review of Ecology, Evolution, and Systematics*, *45*, 247–269. https://doi.org/10.1146/annurev-ecolsys-120213-091935
Passalacqua, P., Belmont, P., Staley, D. M., Simley, J. D., Arrowsmith, J. R., Bode, C. A., et al. (2015). Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: A review. *Earth-Science Reviews*, *148*, 174–193. https://doi.org/10.1016/j.earscirev.2015.05.012
Pennington, D. N., Hansel, J. R., & Gorchov, D. L. (2010). Urbanization and riparian forest woody communities: Diversity, composition, and structure within a metropolitan landscape. *Biological Conservation*, *143*(1), 182–194. https://doi.org/10.1016/j.biocon.2009.10.002
Pinter, N. (2005). One step forward, two steps back on US floodplains. *Science*, *308*, 207. https://doi.org/10.1126/science.1108411
Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., et al. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-020-18321-y
R Core Team. (2020). *R: A language and environment for statistical computing*.
Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929. https://doi.org/10.1111/ecog.02881
Roering, J. J., Mackey, B. H., Marshall, J. A., Sweeney, K. E., Deligne, N. I., Booth, A. M., et al. (2013). 'You are HERE': Connecting the dots with airborne lidar for geomorphic fieldwork. *Geomorphology*, *200*, 172–183. https://doi.org/10.1016/j.geomorph.2013.04.009
Safanelli, J. L., Poppiel, R. R., Ruiz, L. F. C., Bonfatti, B. R., Mello, F. A. D. O., Rizzo, R., & Demattê, J. A. (2020). Terrain analysis in Google Earth engine: A method adapted for high-performance global-scale Analysis. *ISPRS International Journal of Geo-Information*, *9*(6), 400. https://doi.org/10.3390/ijgi9060400
Scheel, K., Morrison, R. R., Annis, A., & Nardi, F. (2019). Understanding the large-scale Influence of levees on floodplain connectivity using a hydrogeomorphic approach. *JAWRA Journal of the American Water Resources Association*, *55*(2), 413–429. https://doi.org/10.1111/1752-1688.12717
Schumm, S. A. (1956). Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey. *The Geological Society of America Bulletin*, *67*(5), 597–646. https://doi.org/10.1130/0016-7606(1956)67[597:eodsas]2.0.co;2
Seaber, P. R., Kapinos, F. P., & Knapp, G. L. (1987). *Hydrologic unit maps: US Geological Survey water supply paper 2294*. US Geological Survey.
Simenstad, C. A., Burke, J. L., O'Connor, J. E., Cannon, C., Heatwole, D. W., Ramirez, M. F., et al. (2011). *Columbia River estuary ecosystem classification: concept and application*.
Sofia, G. (2020). Combining geomorphometry, feature extraction techniques and Earth-surface processes research: The way forward. *Geomorphology*, *355*, 107055. https://doi.org/10.1016/j.geomorph.2020.107055
Sofia, G., Fontana, G. D., & Tarolli, P. (2014). High-resolution topography and anthropogenic feature extraction: Testing geomorphometric parameters in floodplains. *Hydrological Processes*, *28*(4), 2046–2061. https://doi.org/10.1002/hyp.9727
Sparks, R. E., Blodgett, K. D., Casper, A. F., Hagy, H. M., Lemke, M. J., Velho, L. F. M., & Rodrigues, L. C. (2017). Why experiment with success? Opportunities and risks in applying assessment and adaptive management to the Emiquon floodplain restoration project. *Hydrobiologia*, *804*, 177–200. https://doi.org/10.1007/s10750-016-2785-8
Steinfeld, C. M., Kingsford, R. T., & Laffan, S. W. (2013). Semi-automated GIS techniques for detecting floodplain earthworks. *Hydrological Processes*, *27*(4), 579–591. https://doi.org/10.1002/hyp.9244
Steinfeld, C. M. M., & Kingsford, R. T. (2013). Disconnecting the floodplain: Earthworks and their ecological effect on a dryland floodplain in the Murray–Darling basin, Australia. *River Research and Applications*, *29*(2), 206–218. https://doi.org/10.1002/rra.1583
Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B*, *36*(2), 111–133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Strahler, A. N. (1952). Hypsometric (area-altitude) analysis of erosional topography. *The Geological Society of America Bulletin*, *63*(11), 1117–1142. https://doi.org/10.1130/0016-7606(1952)63[1117:haaoet]2.0.co;2

Tobin, G. A. (1995). The levee love affair: A stormy relationship? *JAWRA Journal of the American Water Resources Association*, *31*(3), 359–367. https://doi.org/10.1111/j.1752-1688.1995.tb04025.x

Valentine, A., & Kalnins, L. (2016). An introduction to learning algorithms and potential applications in geomorphometry and Earth surface dynamics. *Earth Surface Dynamics*, *4*(2), 445–460. https://doi.org/10.5194/esurf-4-445-2016

Ward, J. V. (1989). The four-dimensional nature of lotic ecosystems. *Journal of the North American Benthological Society*, *8*(1), 2–8. https://doi.org/10.2307/1467397

Ward, J. V., Tockner, K., & Schiemer, F. (1999). Biodiversity of floodplain river ecosystems: Ecotones and connectivity. *Regulated Rivers: Research & Management*, *15*, 125–139. https://doi.org/10.1002/(sici)1099-1646(199901/06)15:1/3<125::aid-rrr523>3.0.co;2-e

White, G. F., Kates, R. W., & Burton, I. (2001). Knowing better and losing even more: The use of knowledge in hazards management. *Global Environmental Change Part B: Environmental Hazards*, *3*(3), 81–92. https://doi.org/10.3763/ehaz.2001.0308

Wing, O. E., Bates, P. D., Neal, J. C., Sampson, C. C., Smith, A. M., Quinn, N., et al. (2019). A new automated method for improved flood defense representation in large-scale hydraulic models. *Water Resources Research*, *55*(12), 11007–11034. https://doi.org/10.1029/2019wr025957

Wing, O. E., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erickson, T. A. (2017). Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resources Research*, *53*(9), 7968–7986. https://doi.org/10.1002/2017wr020917

Wohl, E. (2005). Disconnected rivers: Human impacts to rivers in the United States. *Reviews in Engineering Geology*, *16*, 19–34.

Wohl, E. (2017). Connectivity in rivers. *Progress in Physical Geography*, *41*(3), 345–362. https://doi.org/10.1177/0309133317714972

Wohl, E. (2018). *Sustaining river ecosystems and water resources*. Springer International Publishing.

Wohl, E., Lininger, K. B., & Baron, J. (2017). Land before water: The relative temporal sequence of human alteration of freshwater ecosystems in the conterminous United States. *Anthropocene*, *18*, 27–46. https://doi.org/10.1016/j.ancene.2017.05.004